Complete Sequencing and Diversity Analysis of the Enterotoxin-Encoding Plasmids in *Clostridium perfringens* Type A Non-Food-Borne Human Gastrointestinal Disease Isolates†

Kazuaki Miyamoto,¹ Derek J. Fisher,^{2,3} Jihong Li,² Sameera Sayeed,² Shigeru Akimoto,¹ and Bruce A. McClane^{2,3*}

Department of Microbiology, Wakayama Medical University School of Medicine, 811-1 Kimiidera, Wakayama, 641-0012, Japan,¹ and Department of Molecular Genetics and Biochemistry² and Molecular Virology and Microbiology Graduate Program,³ University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15261

Received 12 September 2005/Accepted 5 December 2005

Enterotoxin-producing *Clostridium perfringens* type A isolates are an important cause of food poisoning and non-food-borne human gastrointestinal diseases, e.g., sporadic diarrhea (SPOR) and antibiotic-associated diarrhea (AAD). The enterotoxin gene (*cpe*) is usually chromosomal in food poisoning isolates but plasmid-borne in AAD/SPOR isolates. Previous studies determined that type A SPOR isolate F5603 has a plasmid (pCPF5603) carrying *cpe*, IS1151, and the beta2 toxin gene (*cpb2*), while type A SPOR isolate F4969 has a plasmid (pCPF4969) lacking *cpb2* and IS1151 but carrying *cpe* and IS1470-like sequences. By completely sequencing these two *cpe* plasmids, the current study identified pCPF5603 as a 75.3-kb plasmid carrying 73 open reading frames (ORFs) and pCPF4969 as a 70.5-kb plasmid carrying 62 ORFs. These plasmids share an ~35-kb conserved region that potentially encodes virulence factors and carries ORFs found on the conjugative transposon Tn916. The 34.5-kb pCPF4969 variable region contains ORFs that putatively encode two bacteriocins and a two-component regulator similar to VirR/VirS, while the ~43.6-kb pCPF5603 variable region contains a functional *cpb2* gene and several metabolic genes. Diversity studies indicated that other type A plasmid *cpe*⁺/IS1151 SPOR/AAD isolates carry a pCPF5603-like plasmid, while other type A plasmid *cpe*⁺/IS1470-like SPOR/AAD isolates carry a pCPF4969-like plasmid. Tn916-related ORFs similar to those in pCPF4969 (known to transfer conjugatively) were detected in the *cpe* plasmids of other type A SPOR/AAD isolates, as well as in representative *C. perfringens* type B to D isolates carrying other virulence plasmids, possibly suggesting that most or all *C. perfringens* virulence plasmids transfer conjugatively.

Clostridium perfringens is an important pathogen of humans and domestic animals (24, 25). The virulence of this gram-positive bacterium is largely attributable to prolific toxin production; ~15 different *C. perfringens* toxins have been reported in the literature (29). However, individual *C. perfringens* isolates do not produce this entire toxin arsenal, providing pathogenic versatility that allows this bacterium to cause a spectrum of enteric and histotoxic diseases. Isolate-to-isolate toxin production variability also provides the basis for a commonly used classification scheme that assigns *C. perfringens* isolates to one of five toxinotypes (types A to E), based upon their expression of four toxins (α , β , ϵ , and ι toxins) (24). Type A isolates produce α toxin but not β , ϵ , or ι toxin.

Toxin production diversity among *C. perfringens* isolates is probably attributable, at least in part, to the presence of many toxin genes on large plasmids (24). For example, plasmids carry the genes encoding all of the typing toxins except α toxin, as well as the gene (*cpb2*) encoding the recently identified beta2 toxin (CPB2) (24). Despite their obvious pathogenic importance, the large toxin-encoding plasmids of *C. perfringens*

have received limited study (4). To date, only a single *C. perfringens* plasmid (i.e., pCP13) carrying toxin gene sequences has been completely sequenced (27). However, two laboratories recently reported that the pCP13 *cpb2* sequences, which are the only potentially toxin-encoding sequences present on this plasmid, are translationally silent due to the presence of a premature stop codon that prevents CPB2 production (11, 16).

Currently, the best-studied *C. perfringens* virulence plasmids are the enterotoxin-encoding plasmids of type A isolates. Although they represent <5% of all *C. perfringens* isolates, type A isolates producing *C. perfringens* enterotoxin (CPE) are responsible for a very common human food poisoning (17). These bacteria also cause non-food-borne human gastrointestinal (GI) diseases, such as antibiotic-associated diarrhea (AAD) and sporadic diarrhea (SPOR), and (likely) enteric disease in domestic animals (18). In type A food poisoning isolates, the enterotoxin gene (*cpe*) is usually chromosomal (8, 9), whereas type A isolates causing nonfoodborne human GI disease and animal disease isolates typically carry their *cpe* gene on a large plasmid (30). Whether chromosomal or plasmid-borne, the *cpe* gene of type A isolates is immediately preceded by an upstream IS1469 element (6, 22).

Beyond those conserved IS1469-*cpe* sequences, considerable variation exists between the chromosomal and plasmid *cpe* loci of type A isolates. When chromosomal, the IS1469-*cpe* se-

* Corresponding author. Mailing address: E1240 BSTWR, University of Pittsburgh School of Medicine, Pittsburgh, PA 15261. Phone: (412) 648-9022. Fax: (412) 624-1401. E-mail: bamcc@pitt.edu.

† Supplemental material for this article may be found at <http://jb.asm.org/>.

TABLE 1. List of *C. perfringens* isolates used in this study

Isolate group	Strain	Type	<i>cpe</i> location	Date and source	Reference	IS element ^a
Dog	13	A	None		27	
Dysentery/lamb	NCTC8533	B	None	1950s, United Kingdom	http://www.hpa.org.uk/srmd/div_cdmssd_nctc/default.htm	
Struck/sheep	NCTC3182	C	None	1930, unknown		
Sheep	NCTC8346	D	None	1950s, United Kingdom		
Calf	NCTC8084	E	Plasmid ^b	1943, United Kingdom		
Sporadic diarrhea/human	JIR4468 ^c	A	<i>cpe</i> Δ(plasmid)	JIR325(pMRS4969) Cm ^r	5	
	F4969	A	Plasmid	1990's Europe	9	IS1470-like
	F5603	A	Plasmid	1990's Europe		IS1151
	F4013	A	Plasmid	1990's Europe		IS1151
	F4396	A	Plasmid	1990's Europe		IS1470-like
AAD/human	B38	A	Plasmid	1980's, Europe	9	IS1151
	B40	A	Plasmid	1980's, Europe		IS1151
	NB16	A	Plasmid	1980's, Europe		IS1151
	B2	A	Plasmid	1980's, Europe		IS1151
	S43526	A	Plasmid	Late 1990's, NA ^d	30	IS1470-like
	W43181	A	Plasmid	Late 1990's, NA		IS1470-like
	H38094	A	Plasmid	Late 1990's, NA	30	IS1151
	M26413	A	Plasmid	Late 1990's, NA		IS1470-like
	X5722	A	Plasmid	Late 1990's, NA		IS1470-like
	T34058	A	Plasmid	Late 1990's, NA		IS1470-like
Healthy human	MR2-4	A	Plasmid	1990's, Japan	22	IS1470-like

^a Source: Miyamoto et al., 2003.

^b The *cpe* gene associated with type E isolates is silent due to multiple DNA sequence changes.

^c JIR325 is a Rif^r, Nal^r strain 13 derivative transformed with pMRS4969 (pCPR4969 with its *cpe* gene specifically inactivated by insertion of a chloramphenicol resistance determinant).

^d NA, North America.

quences are flanked by two IS1470 sequences, which may represent the boundaries of an integrated ~6.3-kb transposon (6). In contrast, upstream or downstream IS1470 elements are not associated with the plasmid IS1469-*cpe* region of type A isolates; instead, those isolates carry either an IS1151 sequence or a defective IS1470-like sequence downstream of their plasmid *cpe* gene (22, 23). The discovery of two different IS-related sequences downstream of the plasmid *cpe* gene recently led to the identification of at least two distinct *cpe* plasmids in type A isolates causing human non-food-borne GI disease. F5603 (where IS1151 sequences are located downstream of the plasmid *cpe* gene, making this a plasmid *cpe*⁺/IS1151⁺ type A isolate) is a type A SPOR isolate previously reported to carry an ~75-kb plasmid (pCPF5603) with functional *cpe* and *cpb2* genes (22). In contrast, F4969 (where defective IS1470-like sequences are present downstream of the plasmid-borne *cpe* gene, making this a plasmid *cpe*⁺/IS1470-like⁺ type A isolate) is a type A SPOR isolate previously reported to carry an ~75-kb plasmid (pCPF4969) with a functional *cpe* gene but no *cpb2* sequences (22).

Differences between pCPF5603 and pCPF4969 clearly extend beyond the presence or absence of a *cpb2* gene and what IS-related sequence is located downstream of the *cpe* gene. Recent studies (11) indicated that the ~19-kb region between the *cpb2* and *cpe* genes on pCPF5603 is largely conserved among several other plasmid *cpe*⁺/IS1151⁺ type A SPOR or AAD isolates. However, that ~19-kb region is absent from several plasmid *cpe*⁺/IS1470-like⁺ type A SPOR or AAD isolates (including F4969). Currently, it is unclear whether the *cpe* plasmids of *cpe*⁺/IS1151⁺ type A isolates share any similarity beyond the sequenced ~19-kb *cpe*-*cpb2* pCPF5603 region or whether the *cpe* plasmids of *cpe*⁺/IS1151⁺ type A isolates and

cpe⁺/IS1470-like⁺ type A isolates share any similarity outside the sequenced *cpe*-*cpb2* region of pCPF5603.

It has also been shown that pCPF4969 can transfer, via conjugation, between *C. perfringens* isolates (5). Unpublished results referred to in a recent review (24) suggest that at least some type D isolates may also conjugatively transfer the plasmid carrying the epsilon toxin gene (*etx*). Thus, conjugative transfer of *C. perfringens* virulence plasmids, including at least some *cpe*-carrying plasmids, probably contributes to toxin production diversity among *C. perfringens* isolates. Furthermore, Southern blot hybridization studies have detected some similarity between pCPF4969 and pCW3, a *C. perfringens* conjugative plasmid encoding tetracycline resistance (5, 26). However, the transfer mechanism of pCPF4969 or pCW3 is not currently known.

Although the *cpe*-carrying plasmids of type A SPOR/AAD isolates represent a paradigm for *C. perfringens* virulence plasmids, many important questions about these plasmids remain unanswered. For example, do pCPF5603 and pCPF4969 represent the full diversity of *cpe* plasmids among type A SPOR/AAD isolates? Do these two *cpe* plasmids share any common regions? Are other virulence genes present on the unsequenced portions of pCPF5603 and pCPF4969? What mechanism mediates pCPF4969 transfer? Do *cpe* plasmids of other type A AAD/SPOR isolates also conjugatively transfer between *C. perfringens* isolates? If so, do they use transfer mechanisms similar to those of pCPF4969? To gain insights into these important questions, the current study completely sequenced pCPF5603 and pCPF4969 and then used those sequencing results to evaluate, by PCR, the diversity of *cpe* plasmids among type A SPOR/AAD isolates.

MATERIALS AND METHODS

Bacterial culture conditions. *C. perfringens* isolates (Table 1) were grown overnight at 37°C in either TGY (3% tryptic soy broth [TSB; Becton, Dickinson and Company, MD], 2% glucose [Sigma Aldrich Co., MO], 1% yeast extract [Becton, Dickinson and Company, MD], 0.1% L-cysteine [Sigma Aldrich Co., MO]) or FTG (fluid thioglycolate; Difco Laboratories, MI) medium. For construction of plasmid DNA libraries, *C. perfringens* isolates F5603 and F4969 were grown for 6 to 8 h in TGY medium, while JIR4468 (5) was grown for 6 to 8 h in TGY medium with 10 µg/ml chloramphenicol (Sigma Aldrich Co., MO). *Escherichia coli* HB101 (Invitrogen, CA) was grown in TSB (Difco Laboratories, MI). For growth of recombinant *E. coli*, ampicillin (100 to 150 µg/ml; Wako Chemical Co., Japan) was added to either TSB broth or TSB agar plates.

Complete or partial sequencing of the *cpe* plasmids from *C. perfringens* type A SPOR isolates F4969, F5603, and F4013. Using a previously described method (22), crude plasmid extracts were prepared from 1.5-ml cultures of JIR4468 or type A SPOR isolate F4013 or F5603. Those crude plasmid preparations were digested overnight with XbaI or HindIII (Roche Applied Science, IN, or New England Biolabs, MA) and then electrophoresed on a 1% agarose gel. Fragments (4 to 8 kb) were extracted from each gel using the Prep-A Gene DNA purification kit (Bio-Rad, CA) or Quantum Prep Freeze 'N Squeeze DNA gel extraction spin columns (Bio-Rad, CA). The extracted fragments were then ligated into digested and dephosphorylated pBlueScript II SK+ vector (Invitrogen, CA), using T4 DNA ligase (Roche Applied Science, IN) for 16 to 18 h at 16°C. Those recombinant plasmids were transformed into chemically competent *E. coli* HB101 and selected on TSB agar plates containing 100 µg/ml ampicillin.

Preliminary studies (data not shown) revealed that F4969 carries multiple plasmids. Since the presence of several plasmids, of unknown sequence, would technically complicate the isolation and sequencing of pCPF4969 from F4969, we instead performed pCPF4969 sequencing using JIR4468. JIR4468 is a previously prepared (5) *C. perfringens* strain 13 derivative transformed with pMRS4969, which is pCPF4969 with its *cpe* gene specifically inactivated by insertion of a chloramphenicol resistance marker. The rationale for using JIR4468 as the DNA source for pCPF4969 sequencing was that (i) a previous genome sequencing project had identified pCP13 as the only plasmid naturally present in strain 13, (ii) the pCP13 sequence was known, and (iii) 8 kb of sequence containing the complete pCPF4969 *cpe* locus was already available. Using pMRS4969 sequences for primer design, we then completely sequenced pCPF4969.

Furthermore, previous studies had suggested that pCPF4969 carries some sequences homologous to pCW3 (5), so transformants carrying potential *cpe* plasmid sequences were initially identified on the basis of Southern blotting with pCW3 probes. Briefly, transformants carrying DNA from JIR4468 or F5603 were electrophoresed on 1% agarose gels for 16–18 h, plasmid DNA fragments were transferred onto nylon membranes (Roche Applied Science, CA). Those membranes were then subjected to Southern blotting using a randomly digoxigenin (DIG)-labeled plasmid carrying a ClaI fragment of pCW3 (either pJIR15, pJIR16, pJIR18, or pJIR32), which had been prepared as described previously (5). Hybridized probes were detected with CSPD ready-to-use substrate (Roche Applied Science, CA). Transformants carrying plasmid DNA inserts hybridizing with pCW3 probes were then selected and their plasmid inserts were sequenced; the resulting sequence was compared against the previously determined (27) genome sequence of *C. perfringens* strain 13 and (for pCPF4969 sequencing) the pCP13 sequence in order to eliminate transformants carrying non-*cpe* plasmid DNA inserts from the plasmid library.

After assembling large contigs carrying plasmid *cpe* sequences by the above approaches, long-range PCR analysis was performed to close the remaining gaps between the contigs using the Expand Long Template PCR System from Roche Applied Science (CA). Briefly, crude plasmid preparations were used as DNA template for this long-range PCR, and primers 4-21DCMUP/3-03FD0, M-09FD/4-01FD1, and Xb-D2/1-45FD (see Table SI in the supplemental material) were designed to amplify products from the ends of the contigs using the following PCR program: 94°C for 1 min, 35 cycles of 94°C for 20 s, 55°C for 30 s, and 68°C for 8 min, and a final extension of 68°C for 12 min. The obtained long-range PCR products were then used as a DNA sequencing template to close each *cpe* plasmid sequence.

For sequencing the variable region of the *cpe* plasmid (pCPF4013) in type A SPOR isolate F4013, long-range PCR analysis was performed between the conserved region and the *cpb2* gene using *cpb2* gene primers 1F or 2R (11). Direct sequencing was then carried out using these long-range PCR products as a template.

Bioinformatic analyses of *cpe* plasmids. ORFs in pCPF4969, pCPF5603, and the variable region of pCPF4013 were identified using ORF Finder, which is available from NCBI (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>), while puta-

tive Shine-Delgarno sequences were detected manually. Identified ORFs were then used for database searches with FASTA (<http://fasta.genome.jp/>) and BLASTP (<http://www.ncbi.nlm.nih.gov/BLAST/>) to identify homology or similarity with known genes or proteins. Results from these analyses, including all nucleotide sequences and ORFs, are located in GenBank under accession numbers AB236336 (pCPF4969), AB236337 (pCPF5603), and AB236338 (pCPF4013 variable region).

PFGE Southern blotting analysis of *cpe* plasmids in representative *C. perfringens* AAD/SPOR type A isolates. Except for slightly modifying the electrophoresis conditions to improve resolution of *cpe* plasmid size, pulsed-field gel electrophoresis (PFGE) was carried out as previously described by Fisher et al. (11). Briefly, specified AAD or SPOR isolates were grown overnight in TGY medium for DNA plug preparation. Following bacteria lysis, processed plugs were loaded onto 1% PFGE certified agarose gels (Bio-Rad) and run under the following conditions: 6 V/cm using ramped pulse times from 1 to 25 s for 24 h at 14°C. MidRange II PFGE markers (Bio-Rad) were used to determine plasmid size. Following electrophoresis, gels were stained with ethidium bromide to detect DNA. After detection, the DNA was transferred to positively charged nylon membranes (Roche Applied Science), prepared for Southern blotting, and hybridized with a DIG-labeled *cpe* probe as previously described (8). DIG labeling and detection reagents were obtained from Roche Applied Science. CSPD substrate (Roche Applied Science) was used for detection as directed by the manufacturer.

PCR analyses to determine whether *cpe* plasmids of other type A AAD/SPOR isolates carry the conserved region (including putative transfer ORFs) and variable regions of pCPF5603 or pCPF4969. Template DNA for all subsequent PCRs was obtained from colony lysates, which were prepared as described previously (32). Briefly, *C. perfringens* type A SPOR or AAD isolates carrying *cpe* plasmids were grown overnight in FTG broth and then streaked onto brain heart infusion agar plates (Difco Laboratories, MI). Following overnight growth, several colonies were picked from each brain heart infusion agar plate and lysed in sterile water by using a microwave to heat the sample.

Each PCR mixture contained 5 µl of template DNA, 40 µl of TAQ Complete 1.1× Master Mix (Gene Choice, Frederick, Maryland), and 2.5 µl of each primer pair (1 µM final concentration). Primers used for investigating conserved regions are listed in Table SII in the supplemental material, while Tables SIII and SIV list primers for assessing the variable regions of pCPF4969 and pCPF5603, respectively. To amplify portions of the putative pCPF4969 transfer region, the primers listed in Table SV in the supplemental material were used. PCRs were performed in a Techne (Buckhardt, Germany) thermocycler using the following conditions: 94°C for 5 min, 30 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 90 s, and a single extension of 72°C for 5 min for the conserved region; 94°C for 5 min, 40 cycles of 94°C for 1 min, 54°C for 1 min, and 72°C for 100 s, and a single extension of 72°C for 10 min for the pCPF4969 and pCPF5603 variable regions; and 94°C for 5 min, 40 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 100 s, and a single extension of 72°C for 10 min for the transfer region. PCR products were separated on 1% agarose gels and visualized with ethidium bromide staining.

Long-range PCR analyses to determine whether the pCPF4969/pCPF5603 conserved region is present in other *C. perfringens* isolates. The presence and orientation of several genes residing in the conserved region of pCPF4969/pCPF5603 (see Results) was confirmed by long-range PCR analyses of eight type A isolates carrying *cpe* plasmids and four non-type A strains. These long-range PCR analyses used the Expand Long Template PCR system from Roche Applied Science (CA) and the following primers: met-up3FD1/met-up1 for the ORF16 homolog-*dcm* region, ORF13HF/met-up3-4 for the ORF13 homolog-ORF16-homolog region, damF/3-3FU4 for the ORF13 homolog-*dam* region, *cna*-F/dam-F for the *dam*-*cna* region, and 8-118E/C1-6T3D for the *cna*-pCP59 homolog region (primer sequences are listed in Table SVI in the supplemental material). PCR was performed using crude plasmid preparations as a template and the following amplification conditions: 94°C for 1 min, 35 cycles of 94°C for 20 s, 55°C for 30 s, and 68°C for 8 min, and a final extension of 68°C for 12 min.

RESULTS

PFGE Southern blotting analyses to compare *cpe* plasmid size among representative *C. perfringens* type A AAD or SPOR isolates. As mentioned in the introduction, previous studies (22) determined that pCPF4969 and pCPF5603 differ in terms of the IS element downstream of their *cpe* genes, the presence or absence of *cpb2*, and the presence or absence of an ~19-kb

FIG. 1. PFGE/Southern blotting analysis of the *cpe* plasmid in select *C. perfringens* type A isolates. After PFGE, DNA was transferred to nylon membranes and probed with a DIG-labeled *cpe* probe. Southern blotting results are shown for *cpe*⁺/IS1470-like⁺ isolate F4969 (lane 3) and *cpe*⁺/IS1151⁺ isolates F5603 and F4013. Ethidium bromide-stained molecular weight markers are shown in the leftmost lane (sizes in kilobases are listed to the left of the markers).

cpe-cpb2 region (11), i.e., at least two distinct *cpe* plasmids exist among type A non-food-borne human GI disease isolates. In addition, overlapping PCR analyses found variations between the *cpe-cpb2* regions of F5603 and F4013, another plasmid *cpe*⁺/IS1151⁺ type A isolate (11). That preliminary finding indicated that some variation can occur even among *cpe* plasmids carrying the same IS element, although the extent of those variations was unknown.

Collectively, previous studies had only begun exploring the diversity of *cpe* plasmids among type A SPOR and AAD isolates. Even size variations between *cpe* plasmids from different type A SPOR/AAD isolates had not been clearly established due to the limited resolution of prior PFGE/Southern blotting studies, which suggested that pCPF5603 and pCPF4969 were both ~75 kb in size (11). To more accurately determine the size of *cpe* plasmids in type A non-food-borne human GI disease isolates, for the current study we performed PFGE Southern blotting using modified conditions that better resolve ~50- to 100-kb plasmids. When applied to five type A *cpe*⁺/IS1151⁺ SPOR or AAD isolates (F5603, F4013, B2, B38, and H38094) and five type A *cpe*⁺/IS1470-like⁺ SPOR or AAD isolates (F4969, X5722, W43181, S43526, and F4396), these new Southern blotting studies identified three different *cpe* plasmid sizes among these 10 isolates (Fig. 1 shows representative results for type A isolates whose *cpe* plasmids were completely or partially sequenced in this study) (results not shown for other

surveyed isolates). All five surveyed plasmid *cpe*⁺/IS1470-like⁺ isolates were found to carry *cpe* plasmids of ~70 kb. In contrast, the *cpe* plasmid was ~75 kb in all surveyed plasmid *cpe*⁺/IS1151⁺ isolates except for F4013, which carries a *cpe* plasmid of only ~54 kb.

Sequencing of pCPF4969. Previous studies sequenced an ~8-kb region of the conjugative *cpe* plasmid pCPF4969 and found (in order) a putative DNA cytosine methylase (*dcm*), an IS1469 element, the *cpe* gene, and a defective IS1470-like element (22). Coupling those previous findings with the PFGE/Southern blotting results shown in Fig. 1 suggested that ~62 kb of pCPF4969 sequence remained unknown. To determine whether this unknown pCPF4969 sequence could carry other virulence genes or genes potentially involved in conjugation or plasmid replication, pCPF4969 was completely sequenced.

This sequencing analysis indicated that the size of pCPF4969 is 70,480 bp (Fig. 2), in close agreement with the PFGE/Southern blotting results shown in Fig. 1. Using the NCBI ORF Finder, 62 ORFs with an average length of 835 nucleotides were identified, giving pCPF4969 a coding density of 73.5%. Of those 62 ORFs, ~25% classified as encoding hypothetical proteins with unknown functions. The GC content, 26.6%, is similar to that of the two previously sequenced *C. perfringens* plasmids (i.e., the 54.3-kb pCP13 carrying translationally silent *cpb2* sequences and the 10.2-kb, bacteriocin-encoding plasmid pIP404) but slightly lower than that of the *C. perfringens* type A strain 13 genome, 28.6% (27). Interestingly, several ORFs found on pCP13, including PCP56, PCP57 (*cna*), PCP58, PCP59, and PCP12, had similarly arranged homologs on pCPF4969. In addition, pCPF4969 carries homologs to pCP13 ORFs PCP60 and PCP61, which have low similarity at the nucleotide level but strong similarity at the amino acid level.

Known or potential virulence genes carried on pCPF4969 include *cpe*, *srt* (sortase), and *cna* (collagen adhesin). In addition, a putative bacterial two-component regulator, which shows high similarity at the amino acid level (although not at the nucleotide level) to the chromosomal *C. perfringens* VirR/VirS two-component regulator of virulence genes (1), was also identified on pCPF4969. Sequence scanning allowing for up to three base pair mismatches did not reveal the presence of any VirR binding boxes on pCPF4969. Moreover, the *C. perfringens* VirR DNA binding motif SKHR was not identified in the pCPF4969 VirR homolog (20). However, the conserved Fx-HxxKxS VirRc family motif (19) was found in the pCPF4969 VirR homolog.

pCPF4969 also putatively encodes two bacteriocins. The first bacteriocin shows amino acid similarity to propionin SM1 (21) from *Propionibacterium jensenii* (e value, 1e-7) and is encoded by a gene cluster also containing a possible bacteriocin transporter. A second bacteriocin is apparently encoded by an operon comprised of nine ORFs whose products show amino acid similarity to SpaR, -K, -F, -E, -G, -C, -T, -B, and -S (e values ranging from 6e-4 to 3e-116), which is an operon involved in the synthesis and secretion of peptide-derived antimicrobial compounds known as lantibiotics (28).

Concerning putative replication origins, the *ori*-related repeat sequences found on pIP404 (13, 14) were not present on pCPF4969. GC skew analysis has successfully predicted the *ori* region from the *Bacillus thuringiensis* toxin plasmid pBtoxis (2), but similar GC skew analysis performed using Artemis soft-

FIG. 2. Complete or partial maps of the *cpe* plasmids sequenced in this study. Maps display ORFs identified by ORF Finder using the *cpe* plasmid sequences obtained during this study. Complete maps of the *cpe* plasmid are shown for PCPF4969 (A) and PCPF5603 (B), and a partial map is shown for the variable region of PCPF4013 (C). For each map, a blue line highlights the putative variable region, red lines mark the putative conserved region, and an orange line delineates genes comprising the putative transfer region (if shown). Potential roles of each ORF are identified by the key in the lower right-hand corner of the figure.

were failed to identify any putative *ori* region on pCPF4969. ORFs encoding putative primase, resolvase, and helicase were identified from the pCPF4969 sequence.

As mentioned earlier, past studies have shown that pCPF4969 can transfer via conjugation (5). Sequence analysis of pCPF4969 revealed the presence of an ORF cluster encoding proteins with amino acid similarity to those encoded by the conjugative transposon Tn916 from *Enterococcus faecalis* (7). Since Tn916 is capable of transferring in most gram-positive bacteria, the discovery of this Tn916-related conjugation machinery provides an attractive potential mechanism for pCPF4969 conjugative transfer. The putative pCPF4969 transfer region specifically carries ORFs encoding products showing similarity to a hydrolase (homologous to Tn916 ORF14), a muramidase, and an integrase/recombinase as well as Tn916 ORFs 13, 15, 16, 17, and 21. pCPF4969 also appears to encode an FtsK/SpoEIII homolog that shares amino acid similarity with the product of Tn916 ORF21, which is thought to act as a DNA translocase. However, no Tn916 regulatory genes, such as *xis* and *int*, were apparent on pCPF4969. Interestingly, a region downstream of the putative transfer region contains an ORF that putatively encodes a protein whose C-terminal region (amino acids 126 to 223) has homology to the N-terminal half of an integrase. No putative *oriT* sequence homologous with Tn916 *oriT* (15) or with the RS_A sequence of Tn4451 was identified on pCPF4969 (10).

Sequencing of pCPF5603. Previous PCR analyses (11) had determined that the 19-kb *cpb2-cpe* region of pCPF5603 is absent from pCPF4969, while the PFGE/Southern blotting results shown in Fig. 1 indicated that pCPF5603 is ~5 kb larger than pCPF4969. Since those data indicated that substantial size and sequence differences must exist between pCPF4969 and pCPF5603, pCPF5603 was also completely sequenced. Those analyses determined this *cpe* plasmid to be 75,280 bp, with a GC content of 25.3%, similar to the %GC content of pCPF4969 and the two other sequenced *C. perfringens* plasmids (13, 27). pCPF5603 carries 73 ORFs, each with an average length of 819 bp, giving this plasmid a coding density of ~79%. Of the 73 pCPF5603 ORFs, ~25% classified as encoding hypothetical proteins with unknown functions.

Comparison of pCPF5603 and pCPF4969 sequences identified strong similarity across an ~35-kb region extending from *cpe* to the ORF encoding an ATPase of Hsp70 (Fig. 2). This largely conserved sequence includes the putative Tn916-related transfer region and *cpe*, as well as *cna*, *srt*, and the other pCP13-related ORFs present on pCPF4969. The conserved region of pCPF4969 and pCPF5603 contains only two short stretches of divergent sequence (pCPF4969 carries an ~3-kb insertion between *flgJ* and ORF16 and a second 1.9-kb insertion between ORF13 and *fisK4* in the putative transfer region, both of which are missing from pCPF5603).

The conserved region shared by pCPF4969 and pCPF5603 comprises ≤50% of the total sequence for these two *cpe* plasmids, leaving an ~34.8-kb variable region for pCPF4969 and an ~43.6-kb variable region for pCPF5603 (Fig. 2). The size differences between these two variable regions, combined with the total determined sequences for pCPF4969 and pCPF5603, are consistent with PFGE/Southern blotting results shown in Fig. 1, indicating pCPF5603 to be ~5 kb larger than pCPF4969.

Similar to pCPF4969, no obvious replication region was identified in pCPF5603 by either sequence comparison against known *C. perfringens* replication regions or by using Artemis software. Regions of interest that are present on pCPF4969, but absent on pCPF5603, include the VirR/VirS homolog and the putative bacteriocins.

The unique region of pCPF5603 is flanked by two sequences that are, or are related to, insertion sequences, i.e., IS1151 (which lies upstream of *cpe*) and a putative transposase showing similarity (e value, $1e-13$) to a transposase for insertion sequence element ISrm3 from *Bacteroides thetaiotaomicron*. The pCPF5603-specific region carries the potential virulence gene *cpb2*, which is known to be expressed by F5603, as well as a large number of metabolism-related ORFs (11). Of those metabolism-related ORFs, a histidinol phosphatase ORF has been previously found in several *C. perfringens* strains (based on BLAST search results), although not in *C. perfringens* type A strain 13 (27). An ORF carrying the carbohydrate metabolism-related gene, *xrf*, has not been previously reported in *C. perfringens*. Several other potential metabolic ORFs, including those apparently encoding a predicted metal-dependent phosphoesterase, an acyl-CoA synthetase, and a putative acyl carrier protein could play a role in lipid metabolism. An ORF putatively encoding a peptide synthetase was also found. Three ORFs apparently encoding putative ABC type transport system components, two of which showed homology with systems found on the strain 13 chromosome (27), were also identified.

PCR survey to determine whether the pCPF4969/pCPF5603 conserved region ORFs are present in other plasmid *cpe*⁺ *C. perfringens* type A isolates. To survey whether the largely conserved region of pCPF5603/pCPF4969 is present in other type A isolates carrying *cpe* plasmids, two PCR assays were used. The first PCR amplified internal portions of ORFs located in the largely conserved region of the pCPF4969/pCPF5603 plasmids, which allowed us to rapidly assess the presence of those ORFs in other type A isolates carrying *cpe* plasmids. ORFs selected for internal amplification were distributed throughout the common region of pCPF4969 and pCPF5603 and included the Tn916-related ORFs (ORF15, *flgJ*, ORF16, ORF13, and *fisK4*), *dam*, *srt*, *cna*, and the homolog of pCP13 PCP59 (Fig. 3A and Table SVII in the supplemental material). Five plasmid *cpe*⁺/IS1470-like⁺ type A AAD or SPOR isolates (including F4969), one *cpe*⁺/IS1470-like⁺ type A isolate from a healthy human (MR2-4), six plasmid *cpe*⁺/IS1151⁺ type A AAD or SPOR isolates (including F5603), and two *cpe*-negative *C. perfringens* type A strains (strain 13, which carries pCP13, and ATCC 3624) were surveyed with this internal ORF PCR battery.

Consistent with previous sequencing results (27), this PCR battery only amplified the *cna* product from pCP13 (the primers used to amplify internal sequences from *srt* and the pCPF4969/pCPF5603 homolog of PCP59 from pCP13 shared <75% homology with their corresponding pCP13 sequences, explaining why these *srt* and PCP59 primer pairs did not amplify any products from strain 13). Furthermore, no PCR products were amplified from *cpe*-negative type A isolate ATCC 3624 (data not shown). However, this PCR battery produced positive results for all 12 surveyed type A isolates carrying *cpe* plasmids, with the exception that negative results were obtained with (i) isolate MR2-4 using primers for the *fisK4* and

FIG. 3. PCR analyses of the putative conserved region in *cpe* plasmids of type A isolates. (A) The putative conserved regions of pCPF4969 and pCPF5603 are highlighted by gray dashed lines. An initial PCR assay used primers to amplify internal portions of representative genes (shown in bold) to assess the presence of these ORFs among plasmid *cpe*⁺ *C. perfringens* type A isolates. Results for this internal assay were almost always positive for all tested isolates (see Table SVII in the supplemental material). (B) To assess the orientation of ORFs located in the conserved region, primers were designed to link ORFs located throughout the conserved region by long-range PCR. Representative results are shown for *cpe*⁺/IS1470-like⁺ isolates F4969 and MR2-4 and for *cpe*⁺/IS1151⁺ isolates F5603 and F4013. PCR products were run on 0.8% agarose gels and visualized with ethidium bromide. Molecular weight markers and sizes are shown to the left of each gel.

ORF15 homologs and (ii) isolate X5722 using primers for the *ftsK4* homolog (see Table SVII in the supplemental material). Based upon previous PCR results (22), Table SVII also indicates the presence of *dcm* and IS1469 in the conserved region of all these type A plasmid *cpe*⁺/IS1470-like⁺ isolates and type A plasmid *cpe*⁺/IS1151⁺ isolates.

Since the surveyed ORFs in the largely conserved region of pCPF4969/pCPF5603 generally appear to be present in the 12 surveyed plasmid *cpe*⁺ type A human isolates, a second, long-range, PCR assay was used as a rapid screen to determine whether these ORFs are arranged similarly as found in the common region of pCPF4969/pCPF5603, i.e., can these ORFs be connected together in a similar pattern of PCR products as produced from the pCPF5603/pCPF4969 conserved region? Common region ORFs selected for long-range PCR amplification linkage included *dcm*-ORF16, ORF13-ORF16, ORF13-*dam*, *dam*-*cna*, and *cna*-PCP59. Expected product sizes for these PCRs were 7.8, 3.1, 5.0, 5.9, and 2.8 kb from pCPF5603 and 10.4, 2.9, 6.8, 5.6, and 2.7 kb from pCPF4969. The slightly larger size of the *dcm*-ORF16 and ORF13-*dam* PCR products amplified from pCPF4969 versus pCPF5603 is due to the pCPF4969 conserved region carrying (respectively) an ~3-kb insertion between *figJ* and ORF16 and a second ~2-kb insertion containing a 1-kb copy of *ftsKB* inserted between ORF13 and *dam* (Fig. 2). The slightly smaller ORF13-ORF16 PCR

product amplified from pCPF4969 versus pCPF5603 is due to the presence of a 0.2-kb deletion of intervening sequence between these two ORFs in pCPF4969.

When F4969 and F5603 were subjected to this long-range PCR battery, products of the expected size were consistently obtained (Fig. 3B). Three additional plasmid *cpe*⁺/IS1470-like⁺ type A isolates (MR2-4, F4396, and X5722) were also tested with this PCR battery and yielded long-range PCR products consistent with their carrying ORFs in an arrangement similar to that found in the largely conserved region of pCPF4969 (Fig. 3B and data not shown). Likewise, three other plasmid *cpe*⁺/IS1151⁺ type A isolates (F4013, NB16, and H38094) supported amplification of similar-sized PCR products as those amplified from F5603. However, with plasmid *cpe*⁺/IS1151⁺ type A isolate F4013, primers that amplify the *dcm*-ORF16, ORF13-ORF16, and *dam*-*cna* regions amplified PCR products more closely matching the product size expected from the IS1470-like isolate F4969 than from the plasmid *cpe*⁺/IS1151⁺ type A isolate F5603. However, the ORF13-*dam* and *cna*-PCP59 products amplified from F4013 more closely matched the PCR products expected from F5603.

As expected from the pCP13 sequence (27), this long-range PCR battery failed to amplify any products from strain 13, except for the *cna*-PCP59 reaction.

FIG. 4. Overlapping PCR survey of plasmid *cpe*⁺ *C. perfringens* type A isolates using primers designed to amplify the pCPF4969 variable region. (A) Overlapping primers were designed to survey the entire pCPF4969 variable region spanning from the ATPase of Hsp70 class homolog to *cpe*. (B) These 24 reactions were used to survey six plasmid *cpe*⁺/IS1470-like⁺ and six plasmid *cpe*⁺/IS1151⁺ *C. perfringens* type A isolates. (C) Results are shown for two representative *cpe*⁺/IS1470-like⁺ isolates, F4969 and MR2-4. Expected PCR product sizes are shown in panel B below the bars. PCR products were separated on 1% agarose gels and visualized with ethidium bromide staining. Molecular weight markers are indicated at the right of each gel. The gray bars in panel B indicate reactions that were not consistently amplified from all *cpe*⁺/IS1470-like⁺ isolates. Complete results from this PCR survey are shown in Table SVIII in the supplemental material.

Sequencing of the variable region in the F4013 *cpe* plasmid.

In addition to the minor atypical features within its conserved region identified in Fig. 3, two other results indicated that the *cpe* plasmid of F4013 is unusual. A previous overlapping PCR assay covering the 19-kb *cpe*-*cpb2* variable region of pCPF5603 had identified the variable region of pCPF4013 as being different from other type A *cpe*⁺/IS1151⁺ plasmids (11). Second, the PFGE/Southern blotting results shown in Fig. 1 had demonstrated that pCPF4013 is much smaller than either pCPF4969 or pCPF5603. Collectively, these results suggested that the smaller pCPF4013 could be carrying a third kind of *cpe* plasmid variable region. To assess this possibility, the pCPF4013 variable region was completely sequenced.

This sequencing analysis confirmed significant differences between the variable region of pCPF4013 versus pCPF4969 or pCPF5603. The pCPF4013 variable region is significantly smaller (~24 kb) than the variable region of either pCPF4969 or pCPF5603, although it more closely resembles the variable region of pCPF5603 (Fig. 2). Most differences between pCPF4013 and pCPF5603 can be attributed to ~20 kb of the pCPF5603 variable region being absent from pCPF4013. Specifically, pCPF4013 lacks pCPF5603 variable region sequences present between the ORF putatively encoding an ATPase of Hsp70 and the *padR* ORF, i.e., nearly the entire ~20-kb metabolism-related gene cluster of pCPF5603 is absent from

pCPF4013. Minor differences between pCPF5603 and pCPF4013, including the absence of *thiF* and *comEC* from pCPF4013, were also apparent in the region between *cpb2* and *cpe*.

Overlapping PCR analyses to determine whether the pCPF4969 variable region is present among other type A isolates carrying *cpe* plasmids. The presence of the pCPF4969 variable region among type A isolates carrying *cpe* plasmids was determined using an overlapping PCR battery that included 24 pairs of overlapping primers to directly connect ~36 kb of pCPF4969 sequence that spans from *cpe* to the ATPase of the Hsp70 class homolog (Fig. 4A and B). When used with F4969 as a positive control, this PCR battery produced products of the expected size for all 24 PCRs (Fig. 4C and Table SVIII in the supplemental material). When similarly applied to five other *cpe*⁺/IS1470-like⁺ type A isolates (four AAD or SPOR isolates and one isolate from a healthy human), this PCR battery produced results consistent with these five isolates all possessing the entire pCPF4969 variable region (Fig. 4C shows representative PCR results; also see Table SVIII in the supplemental material). The only consistent exceptions were that reactions 20 and 21 (Fig. 4B), which pass through a putative transposon gene in pCPF4969, were not amplified from four of the other five surveyed plasmid *cpe*⁺/IS1470-like⁺ type A isolates.

In contrast, six plasmid *cpe*⁺/IS1151⁺ *C. perfringens* type A

A) pCPF5603 variable region map and PCRs

B) PCR Results

FIG. 5. Analysis of the pCPF5603 variable region in plasmid *cpe*⁺ *C. perfringens* type A isolates using an overlapping PCR assay. (A) The pCPF5603 variable region (from the ATPase of Hsp70 class homolog to *cpe*) was amplified using 33 overlapping PCRs (black bars). PCR products were analyzed on 1% agarose gels and visualized with ethidium bromide staining. Using this PCR assay, the presence of the pCPF5603 variable region was assessed in six plasmid *cpe*⁺/IS1470-like⁺ and six plasmid *cpe*⁺/IS1151⁺ *C. perfringens* type A isolates. PCR reaction 7 (gray bar) was not consistently amplified by *cpe*⁺/IS1151⁺ isolates. (B) Representative PCR survey results are shown for *cpe*⁺/IS1151⁺ isolates F5603 and B40. The locations of molecular weight markers are indicated to the right of each gel. PCR results for all surveyed isolates are listed in Table SIX in the supplemental material.

isolates (all associated with AAD or SPOR) failed to support amplification of nearly all 24 reactions included in the pCPF4969 variable region PCR battery. The only consistent exceptions were for PCRs 4, 9, and 24, which amplified a region encoding two hypothetical proteins, a region encoding a homolog to PCP12 from pCP13 and a hypothetical protein, and a region encoding an internal portion of the Hsp70 ATPase homolog, respectively. All three of these DNA sequences should yield positive PCR results from *cpe*⁺/IS1151⁺ type A isolates since they are also found on pCPF5603. Supporting the specificity of this PCR battery, strain 13 did not yield positive PCRs for the pCPF4969 variable region, consistent with the complete absence of this region from pCP13.

Overlapping PCR analyses to determine whether the pCPF5603 variable region is present among other type A isolates carrying *cpe* plasmids. A second overlapping PCR assay was designed to survey the presence of the pCPF5603 variable region among type A isolates carrying *cpe* plasmids. This PCR battery included 33 overlapping PCRs that span the entire ~45-kb pCPF5603 region from *cpe* to the Hsp70 ATPase homolog (Fig. 5A). When applied to F5603 as a positive control, this PCR battery amplified products of the expected size for all 33 reactions (Fig. 5B and Table SIX in the supplemental material). With minor exceptions, this PCR battery also consistently amplified products of the expected size from four of five other surveyed plasmid *cpe*⁺/IS1151⁺ type A AAD or SPOR isolates (representative results are shown in Fig. 5B, with all

FIG. 6. PCR survey for the presence of the Tn916-related putative transfer region of pCPF4969 or pCPF5603 in other plasmid *cpe*⁺ *C. perfringens* type A isolates. (A) Primers were designed to amplify a portion (dashed line) of the Tn916-related putative transfer regions from either pCPF4969 or pCPF5603. (B) Thirteen or eleven overlapping PCRs were used to survey six plasmid *cpe*⁺/IS1470-like⁺ and six plasmid *cpe*⁺/IS1151⁺ *C. perfringens* type A isolates for the presence of the putative transfer region. Reactions 1 to 10 were identical for both assays, whereas reactions 11A to 13A and reaction 11B were designed to amplify specific sequences in the pCPF4969 or pCPF5603 putative transfer region, respectively, as shown in panel B. Expected PCR product sizes are listed below the black bars in panel B. (C) Representative PCR results are shown for F5603 and F4969. Complete results from this survey are listed in Table SX in the supplemental material. PCR products were run on 1% agarose gels and stained with ethidium bromide. The positions of molecular weight markers are shown to the right of each gel.

results summarized in Table SIX in the supplemental material). The exception was *cpe*⁺/IS1151⁺ type A SPOR isolate F4013, which failed to amplify a substantial number of pCPF5603 variable region genes, consistent with the deletion of 20 kb of sequence from the pCPF4013 variable region detected by sequencing (Fig. 2).

In contrast, when this same PCR battery was applied to six plasmid *cpe*⁺/IS1470-like⁺ type A isolates (five associated with AAD or SPOR and one from a healthy human), only a few primer pairs consistently amplified products of the expected size. Those primers amplified the PCP12 homolog hypothetical protein (reaction 16) and regions flanking two hypothetical proteins (reactions 27 and 28), all of which involve sequences common to both the pCPF5603 and pCPF4969 sequence (Fig. 2).

As a negative control, this PCR battery did not amplify any products of the expected size from strain 13, which is consistent with the absence of this region from pCP13 (27).

Overlapping PCR analyses to determine whether the putative Tn916-related transfer region is present among other type A isolates carrying *cpe* plasmids. Initial long-range PCR testing for the presence of the pCPF4969 and pCPF5603 conserved regions (Fig. 3) indicated that all surveyed type A human isolates carrying *cpe* plasmids have the Tn916-related region found in the conjugative plasmid pCPF4969. However,

those initial experiments sampled only a limited number of type A isolates carrying *cpe* plasmids and did not directly connect the Tn916-related region to the *cpe* gene present in those other type A isolates, i.e., conceivably, some or all of those isolates could carry their Tn916-related genes on either another, non-*cpe* carrying, plasmid or on their chromosome. Since establishing the presence of Tn916-related genes on *cpe* plasmids in other AAD/SPOR type A isolates has potential implications for predicting the conjugative transferability of those *cpe* plasmids, an overlapping PCR battery was constructed. This assay, which directly links the plasmid-borne *cpe* gene to the Tn916-related genes on pCPF4969 and pCPF5603, consisted of 10 common reactions and either 11 or 13 total PCRs (Fig. 6A and B), depending upon whether the primers were designed to amplify the transfer region of pCPF5603 or pCPF4969 (which contains a small insertion between the *figJ* ORF and Tn916 ORF16).

This transfer region PCR battery consistently amplified products of the expected size from all 12 surveyed type A isolates carrying *cpe* plasmids, 11 of which were associated with AAD or SPOR. The pattern of reactions amplifying products was highly consistent (see Table SX in the supplemental material; representative results are shown in Fig. 6C), depending upon whether the type A isolate was a plasmid *cpe*⁺/IS1151⁺ isolate (like F5603) or a plasmid *cpe*⁺/IS1470-like⁺ isolate

FIG. 7. PCR survey for the presence of the *cpe* plasmid conserved region in *C. perfringens* type B to D isolates. (A) Using the long-range PCR assay shown in Fig. 3, *C. perfringens* type B to D isolates were analyzed for the presence of the conserved region. (B) PCR products were separated on 0.8% agarose gels and stained with ethidium bromide. Results for each PCR are shown. Molecular weight markers are shown to the left of each gel. Isolate names and types are at the top of each gel.

(like F4969). The major exception was that reaction 7, which uses primers to connect the hydrolase homolog to Tn916 ORF15, did not amplify a product in several isolates, regardless of their IS class (*cpe*⁺/IS1151⁺ or *cpe*⁺/IS1470-like⁺). Consistent with the long-range PCR results shown in Fig. 3, the plasmid *cpe*⁺/IS1151⁺ type A isolate F4013 was negative for pCPF5603-based PCR 11B but positive for pCPF4969-based reactions 11A to 13A.

Long-range PCR to determine whether the conserved region, including the putative pCPF4969 transfer ORFs, is present in other *C. perfringens* types. *C. perfringens* type B to E isolates all carry large virulence plasmids, which carry both the *etx* gene and the beta toxin gene (*cpb*), *cpb* only, *etx* only, or the iota toxin genes only, respectively (24). In addition, unpublished reports suggest that, like pCPF4969, the *etx* plasmid of type D isolates may transfer conjugatively. Collectively, those facts raised interest in testing whether the Tn916-related putative transfer ORFs of pCPF4969 and *cpe* plasmids in type A isolates are also present in representative *C. perfringens* type B to E isolates. To address that question, the long-range PCRs shown in Fig. 7A were used. Results obtained indicated the

presence of the putative transfer region ORFs in each of the representative type B to E isolates (Fig. 7B). Each surveyed type B to E isolate supported PCR amplification of all five reactions that detect the presence of conserved region genes (including transfer ORFs). PCR product sizes generally matched those amplified from plasmid *cpe*⁺ *C. perfringens* type A isolates. However, small variations in PCR product sizes for the *cna*-PCP59, Tn916 ORF13-*dam*, and *dcm*-Tn916 ORF16 PCRs were noted for particular isolates (Fig. 7B).

DISCUSSION

Toxin-encoding plasmids are important for the pathogenesis of many *C. perfringens* infections, but they have been poorly studied. The large size and low copy number of these plasmids complicates their isolation, analysis, and sequencing, as does the fact that many individual *C. perfringens* isolates (including both F4969 and F5603 [unpublished observations by Miyamoto et al.]) carry multiple plasmids. We overcame these technical challenges with a unique sequencing strategy that (i) employed JIR4468 (5), a strain 13 derivative transformed with

pMRS4969 (a slightly modified pCPF4969) to help with initial sequencing of pCPF4969, thereby eliminating other unsequenced plasmids from our screening library; and (ii) exploited the partial homology (5) between pCPF4969 and the *C. perfringens* tetracycline resistance-encoding plasmid pCW3 (26) by using pCW3-based probes for library screening, which substantially enriched for transformants carrying *cpe* plasmid inserts. Long-range PCR then closed gaps between contigs assembled from this pCW3 probe screening, allowing complete sequencing of the first two *C. perfringens* virulence plasmids carrying one or more functional toxin genes.

Our sequencing results revealed that pCPF4969 and pCPF5603 share an ~35-kb conserved region, along with (respectively) ~35- or ~40-kb variable regions. PCR surveys of other AAD or SPOR type A isolates from varied geographic origins strongly suggest that most *cpe*⁺/IS1151⁺ plasmids in AAD/SPOR type A isolates are similar to pCPF5603. The smaller F4013 *cpe* plasmid simply appears to be a pCPF5603 variant with an ~20-kb deletion in its variable region. Similarly, PCR surveys suggest that the *cpe*⁺/IS1470-like⁺ plasmids of most other AAD/SPOR type A isolates resemble pCPF4969. The existence of just two major *cpe* plasmid families in type A isolates is further supported by PCR results detecting a pCPF4969-like *cpe* plasmid in two type A plasmid *cpe*⁺/IS1470-like animal disease isolates (22 and data not shown).

The two major *cpe* plasmid families of type A isolates apparently share a common evolutionary origin, based upon the strong similarity of the pCPF4969 and pCPF5603 conserved regions. This similarity includes the presence of nearly identical Tn916-related sequences, which suggests that an early step in *cpe* plasmid evolution involved Tn916 integration onto a *C. perfringens* plasmid, perhaps one related to pCP13 (given the presence of some pCP13 ORF homologues in the pCPF5603/pCPF4969 conserved region). This integration likely created a progenitor conjugative plasmid that has subsequently transferred among type A isolates.

In one recipient, the progenitor plasmid probably acquired an IS1469-*cpe*-IS1470 element to form the pCPF4969 *cpe* plasmid family. This progenitor plasmid may have simultaneously acquired both the IS1469-*cpe*-IS1470-like locus and the pCPF4969 variable region, consistent with the ability of IS1469 and IS1470 to mediate DNA excision (6). Alternatively, since phage-related genes are present immediately downstream of the IS1470-like sequences of pCPF4969, phage DNA carrying the pCPF4969 variable region may have integrated near the *dcm* region of the progenitor plasmid, followed by later insertion of an IS1469-*cpe*-IS1470 element near *dcm*. This phage scenario offers interspecies DNA transfer as a potential explanation for the pCPF4969 variable region containing DNA sequences (e.g., the *spa* genes) also found in other gram-positive bacteria. After IS1470 localized in pCPF4969, a mutation probably created the defective IS1470-like sequence found in the current pCPF4969.

Similarly, pCPF5603 may have resulted from (i) IS1151, IS1469, or the putative ISRM3-like transposase mediating simultaneous insertion of *cpe* and the pCPF5603 variable region into the progenitor conjugative plasmid or (ii) initial integration of some genetic element containing most pCPF5603 variable region sequences near *dcm*, followed by later insertion of

an IS1151-*cpe*-IS1469 element, also near *dcm*. The *C. perfringens* chromosome may have been one source for pCPF5603 variable region DNA, given the presence of many metabolic genes (including some found on the *C. perfringens* strain 13 chromosome) in this region.

The presence of both the IS1469-*cpe*-IS1151 and the IS1469-*cpe*-IS1470 regions near *dcm* on pCPF5603 and pCPF4969 is notable. This association is consistent with the *dcm* region of these plasmids representing a hot spot for IS element (and perhaps phage) insertion. Since *dcm* is also present in type B to D isolates (22), it will be interesting to determine whether toxin genes present on type B to D virulence plasmids are also located near *dcm*.

The same progenitor conjugative plasmid putatively involved in pCPF4969/pCPF5603 evolution may also have served as the backbone for other *C. perfringens* virulence plasmids. Consistent with this possibility are our PCR results demonstrating that the pCPF4969/pCPF5603 Tn916-related region (and some other conserved region ORFs) is present in four representative type B to E isolates carrying virulence plasmids encoding epsilon, beta, or iota toxins, while those sequences are absent from two type A strains (strain 13 and ATCC 3624) lacking plasmids that carry functional toxin genes. The Tn916-related region in the surveyed type B to D isolates is not present on a *cpe* plasmid since those isolates are *cpe* negative. Consistent with *C. perfringens* virulence plasmids sharing a common origin, it has already been proposed that type E plasmids arose from integration of a mobile genetic element carrying iota toxin genes onto an IS1151⁺/*cpe*⁺ plasmid in a type A isolate (3). Sequencing of type B to E virulence plasmids is under way to conclusively assess the similarity between other *C. perfringens* toxin-encoding plasmids and pCPF4969/pCPF5603.

Tn916 can conjugatively transfer among most gram-positive bacteria and can mobilize plasmids into which it has inserted (7). Therefore, the presence of Tn916-related ORFs in pCPF4969 represents an attractive explanation for the conjugative transfer of this *cpe* plasmid (5). If Tn916-related ORFs mediate pCPF4969 conjugation, the presence of similar Tn916 ORFs in all other surveyed type A AAD/SPOR isolates carrying *cpe* plasmids predicts that most or all *cpe* plasmids are transferable by conjugation. By further extension, detection of Tn916 ORFs in *C. perfringens* type B to E isolates raises the possibility of other *C. perfringens* toxin-encoding plasmids also being conjugative, consistent with unpublished results purportedly demonstrating conjugative transfer of an epsilon toxin-encoding plasmid (24). If confirmed, conjugative transfer of type B to E virulence plasmids may (as previously proposed for pCPF4969 [5]) contribute to enteric disease by allowing a small initial infecting dose of a type B to E isolate to transfer its virulence plasmid to the *C. perfringens* type A isolates present in the normal flora, thereby amplifying the infection.

If the Tn916-related region mediates conjugative transfer of type A *cpe* plasmids and type B to E virulence plasmids, it is notable that some Tn916 genes are missing from pCPF5603/pCPF4969 and those Tn916 genes that are present have a significantly rearranged gene order (Fig. 8). Nevertheless, both *cpe* plasmid families carry the Tn916 ORF 13–17 region involved in conjugative transfer (7). pCPF4969/pCPF5603 also carry an FtsK/SpoEIII homolog ORF with amino acid similar-

FIG. 8. Comparison of the pCPF5603/pCPF4969 putative transfer region gene order and the gene order of the conjugative transposon Tn916. For comparison, the putative transfer regions of pCPF5603 and pCPF4969 are shown schematically above the Tn916 genome. ORFs important in Tn916 conjugative transfer are highlighted by the shaded bar beneath the Tn916 schematic. The bar increases in darkness to indicate ORFs known to be required for conjugative transfer of Tn916.

ity to Tn916 ORF21, which encodes a DNA translocase. This FtsK/SpoEIII homolog may serve two functions for *cpe* plasmids in type A isolates, i.e., translocating these plasmids into the forespore during sporulation (thus ensuring stable inheritance of the *cpe* plasmid in germinating cells) and transferring DNA from donor to recipient cells during conjugation.

Both the pCPF4969 and pCPF5603 *cpe* plasmid families carry additional ORFs (besides *cpe*) that encode recognized or potential virulence factors. Like pCP13, these *cpe* plasmids carry a *cna* homolog that could encode a collagen adhesin (27). These *cpe* plasmids also carry a *srt* homolog that could encode a sortase; in several gram-positive pathogens, sortases anchor surface virulence factors (33). The pCPF4969 *cpe* plasmid family also carries ORFs with similarity, at the amino acid level, to the *C. perfringens* VirR/VirS two-component regulator that positively regulates chromosomal virulence genes encoding alpha toxin and perfringolysin O (1, 25). While lacking VirR/VirS homologs, pCPF5603 carries a functional *cpb2* gene that could be important for virulence since CPB2 can damage CaCo-2 cultured human colonic cells, i.e., CPB2 and CPE may cocontribute to intestinal disease caused by AAD/SPOR isolates carrying *cpe* plasmids of the pCPF5603 family (11).

The pCPF4969 *cpe* plasmid family also apparently encodes the bacteriocin propionin SM1 and a lantibiotic bacteriocin encoded by the *spa* operon. Another sequenced *C. perfringens* plasmid, pIP404 (which carries no virulence genes or Tn916-related genes) also encodes a bacteriocin, Bcn5 (12). While sequencing F5603, we found a second large plasmid (besides pCPF5603) also encoding Bcn5 (Miyamoto et al., unpublished

observation). The common presence of bacteriocin genes on sequenced *C. perfringens* plasmids suggests that bacteriocin production may be beneficial to *C. perfringens*, perhaps by reducing competition from other bacteria when *C. perfringens* is present in the soil or GI tract. The presence of pCPF4969

FIG. 9. Gene order comparison of the *spa* operon from *Bacillus subtilis* ATCC 6633 and the putative *spa* operon from *C. perfringens* pCPF4969. The putative *spa* operon found on pCPF4969 is shown below the *spa* operon identified in *B. subtilis* ATCC 6633 to compare the two gene arrangements.

bacteriocin-encoding ORFs also explains difficulties in using F4969 for mating experiments to demonstrate pCPF4969 conjugative transfer (5). Finally, it is notable that, relative to the *spa* operon of other gram-positive bacteria (31), the bacteriocin-related *spa* ORFs on pCPF4969 are atypically arranged and are missing some ORFs (Fig. 9).

While the current studies offer new insights into the sequences and diversity of the *cpe* plasmids in type A isolates, additional research is needed to identify sequences involved in plasmid replication and to confirm the putative role of Tn916-related genes in conjugative transfer. Other studies should evaluate whether the putative virulence genes identified on these *cpe* plasmids actually contribute to enteric pathogenesis. Finally, sequence comparisons of the *cpe* plasmids of type A isolates against the virulence plasmids of *C. perfringens* type B to E isolates will reveal whether these virulence plasmids share a common origin.

ACKNOWLEDGMENTS

This research was supported by grants R37 AI19844-23, RO1 AI056177-03, T32 AI49820, and T32 AI060525 from the National Institute of Allergy and Infectious Diseases of the U.S. Public Health Service.

We thank Mieko Tamura for technical assistance and Julian Rood for providing the *E. coli* transformants carrying pCW3 fragments.

REFERENCES

1. Ba-Thein, W., M. Lyrstis, K. Ohtani, I. T. Nisbet, H. Hayashi, J. I. Rood, and T. Shimizu. 1996. The *virR/virS* locus regulates the transcription of genes encoding extracellular toxin production in *Clostridium perfringens*. *J. Bacteriol.* **178**:2514–2520.
2. Berry, C., S. O'Neil, E. Ben-Dov, A. F. Jones, L. Murphy, M. A. Quail, M. T. Holden, D. Harris, A. Zaritsky, and J. Parkhill. 2002. Complete sequence and organization of pBtoxis, the toxin-coding plasmid of *Bacillus thuringiensis* subsp. *israelensis*. *Appl. Environ. Microbiol.* **68**:5082–5095.
3. Billington, S. J., E. U. Wiecek, M. R. Sarker, D. Bueschel, J. G. Songer, and B. A. McClane. 1998. *Clostridium perfringens* type E animal enteritis isolates with highly conserved, silent enterotoxin gene sequences. *Infect. Immun.* **66**:4531–4536.
4. Bruggemann, H. 2005. Genomics of clostridial pathogens: implication of extrachromosomal elements in pathogenicity. *Curr. Opin. Microbiol.* **8**:601–605.
5. Brynestad, S., M. R. Sarker, B. A. McClane, P. E. Granum, and J. I. Rood. 2001. Enterotoxin plasmid from *Clostridium perfringens* is conjugative. *Infect. Immun.* **69**:3483–3487.
6. Brynestad, S., B. Synstad, and P. E. Granum. 1997. The *Clostridium perfringens* enterotoxin gene is on a transposable element in type A human food poisoning strains. *Microbiology* **143**:2109–2115.
7. Clewell, D. B., S. E. Flannagan, and D. D. Jaworski. 1995. Unconstrained bacterial promiscuity: the Tn916-Tn1545 family of conjugative transposons. *Trends Microbiol.* **3**:229–236.
8. Collic, R. E., and B. A. McClane. 1998. Evidence that the enterotoxin gene can be episomal in *Clostridium perfringens* isolates associated with non-food-borne human gastrointestinal diseases. *J. Clin. Microbiol.* **36**:30–36.
9. Cornillot, E., B. Saint-Joanis, G. Daube, S. Katayama, P. E. Granum, B. Canard, and S. T. Cole. 1995. The enterotoxin gene (*cpe*) of *Clostridium perfringens* can be chromosomal or plasmid-borne. *Mol. Microbiol.* **15**:639–647.
10. Crellin, P. K., and J. I. Rood. 1998. Tn4451 from *Clostridium perfringens* is a mobilizable transposon that encodes the functional Mob protein, TnpZ. *Mol. Microbiol.* **27**:631–642.
11. Fisher, D. J., K. Miyamoto, B. Harrison, S. Akimoto, M. R. Sarker, and B. A. McClane. 2005. Association of beta2 toxin production with *Clostridium perfringens* type A human gastrointestinal disease isolates carrying a plasmid enterotoxin gene. *Mol. Microbiol.* **56**:747–762.
12. Garnier, T., and S. T. Cole. 1986. Characterization of a bacteriocinogenic plasmid from *Clostridium perfringens* and molecular genetic analysis of the bacteriocin-encoding gene. *J. Bacteriol.* **168**:1189–1196.
13. Garnier, T., and S. T. Cole. 1988. Complete nucleotide sequence and genetic organization of the bacteriocinogenic plasmid, pIP404, from *Clostridium perfringens*. *Plasmid* **19**:134–150.
14. Garnier, T., and S. T. Cole. 1988. Identification and molecular genetic analysis of replication functions of the bacteriocinogenic plasmid pIP404 from *Clostridium perfringens*. *Plasmid* **19**:151–160.
15. Jaworski, D. D., and D. B. Clewell. 1995. A functional origin of transfer (*oriT*) on the conjugative transposon Tn916. *J. Bacteriol.* **177**:6644–6651.
16. Jost, B. H., S. J. Billington, H. T. Trinh, D. M. Bueschel, and J. G. Songer. 2005. Atypical *cpb2* genes, encoding beta2-toxin in *Clostridium perfringens* isolates of nonporcine origin. *Infect. Immun.* **73**:652–656.
17. McClane, B. A. 1997. *Clostridium perfringens*, p. 305–326. In M. P. Doyle, Larry R. Beuchat, and Thomas J. Montville (ed.), *Food microbiology: fundamentals and frontiers*. ASM Press, Washington, D.C.
18. McClane, B. A., David M. Lyster, J. Scott Moncrief, and Tracy D. Wilkins. 2000. Enterotoxigenic clostridia: *Clostridium perfringens* type A and *Clostridium difficile*, p. 551–562. In V. A. Fischetti, Richard P. Novick, Joseph J. Ferretti, Daniel A. Portnoy, and Julian I. Rood (ed.), *Gram-positive pathogens*. ASM Press, Washington, D.C.
19. McGowan, S., I. S. Lucet, J. K. Cheung, M. M. Awad, J. C. Whisstock, and J. I. Rood. 2002. The FxRxHrS motif: a conserved region essential for DNA binding of the VirR response regulator from *Clostridium perfringens*. *J. Mol. Biol.* **322**:997–1011.
20. McGowan, S., J. R. O'Connor, J. K. Cheung, and J. I. Rood. 2003. The SKHR motif is required for biological function of the VirR response regulator from *Clostridium perfringens*. *J. Bacteriol.* **185**:6205–6208.
21. Miescher, S., M. P. Stierli, M. Teuber, and L. Meile. 2000. Propionin SM1, a bacteriocin from *Propionibacterium jensenii* DF1: isolation and characterization of the protein and its gene. *Syst. Appl. Microbiol.* **23**:174–184.
22. Miyamoto, K., G. Chakrabarti, Y. Morino, and B. A. McClane. 2002. Organization of the plasmid *cpe* locus in *Clostridium perfringens* type A isolates. *Infect. Immun.* **70**:4261–4272.
23. Miyamoto, K., Q. Wen, and B. A. McClane. 2004. Multiplex PCR genotyping assay that distinguishes between isolates of *Clostridium perfringens* type A carrying a chromosomal enterotoxin gene (*cpe*) locus, a plasmid *cpe* locus with an IS1470-like sequence, or a plasmid *cpe* locus with an IS1151 sequence. *J. Clin. Microbiol.* **42**:1552–1558.
24. Petit, L., M. Gibert, and M. R. Popoff. 1999. *Clostridium perfringens*: toxinotype and genotype. *Trends Microbiol.* **7**:104–110.
25. Rood, J. I. 1998. Virulence genes of *Clostridium perfringens*. *Annu. Rev. Microbiol.* **52**:333–360.
26. Rood, J. I., V. N. Scott, and C. L. Duncan. 1978. Identification of a transferable tetracycline resistance plasmid (pCW3) from *Clostridium perfringens*. *Plasmid* **1**:563–570.
27. Shimizu, T., K. Ohtani, H. Hirakawa, K. Ohshima, A. Yamashita, T. Shiba, N. Ogasawara, M. Hattori, S. Kuhara, and H. Hayashi. 2002. Complete genome sequence of *Clostridium perfringens*, an anaerobic flesh-eater. *Proc. Natl. Acad. Sci. USA*. **99**:996–1001.
28. Siezen, R. J., O. P. Kuipers, and W. M. de Vos. 1996. Comparison of lantibiotic gene clusters and encoded proteins. *Antonie Leeuwenhoek* **69**:171–184.
29. Smedley, J. G., III, D. J. Fisher, S. Sayeed, G. Chakrabarti, and B. A. McClane. 2004. The enteric toxins of *Clostridium perfringens*. *Rev. Physiol. Biochem. Pharmacol.* **152**:183–204.
30. Sparks, S. G., R. J. Carman, M. R. Sarker, and B. A. McClane. 2001. Genotyping of enterotoxigenic *Clostridium perfringens* fecal isolates associated with antibiotic-associated diarrhea and food poisoning in North America. *J. Clin. Microbiol.* **39**:883–888.
31. Twomey, D., R. P. Ross, M. Ryan, B. Meaney, and C. Hill. 2002. Lantibiotics produced by lactic acid bacteria: structure, function and applications. *Antonie Leeuwenhoek* **82**:165–185.
32. Wen, Q., K. Miyamoto, and B. A. McClane. 2003. Development of a duplex PCR genotyping assay for distinguishing *Clostridium perfringens* type A isolates carrying chromosomal enterotoxin (*cpe*) genes from those carrying plasmid-borne enterotoxin (*cpe*) genes. *J. Clin. Microbiol.* **41**:1494–1498.
33. Zink, S. D., and D. L. Burns. 2005. Importance of *srtA* and *srtB* for growth of *Bacillus anthracis* in macrophages. *Infect. Immun.* **73**:5222–5228.